



THE UNIVERSITY *of* EDINBURGH

## Edinburgh Research Explorer

### Mellow babies and mellow toddlers

**Citation for published version:**

Levi, D, Ibrahim, R, Malcolm, R & Macbeth, A 2018, 'Mellow babies and mellow toddlers: Effects on maternal mental health of a group-based parenting intervention for at-risk families with young children', *Journal of Affective Disorders (JAD)*, vol. 246, pp. 820-827. <https://doi.org/10.1016/j.jad.2018.12.120>

**Digital Object Identifier (DOI):**

[10.1016/j.jad.2018.12.120](https://doi.org/10.1016/j.jad.2018.12.120)

**Link:**

[Link to publication record in Edinburgh Research Explorer](#)

**Document Version:**

Peer reviewed version

**Published In:**

Journal of Affective Disorders (JAD)

**General rights**

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

**Take down policy**

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact [openaccess@ed.ac.uk](mailto:openaccess@ed.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.



**Mellow Babies and Mellow Toddlers: Effects on maternal mental health of a group-based parenting intervention for at-risk families with young children**

Dalia Levi MSc<sup>1</sup>, Raquib Ibrahim MSc<sup>2</sup>; Ruairidh Malcolm MSc<sup>2</sup>, Angus MacBeth PhD<sup>1\*</sup>.

<sup>1</sup>University of Edinburgh, Scotland, UK.

<sup>2</sup> Mellow Parenting, Scotland, UK.

\*Corresponding Author: Dr Angus MacBeth

School of Health in Social Science

The University of Edinburgh

Rm 3.06A

Doorway 6, Medical Quad

Teviot Place

Edinburgh, EH8 9AG

Tel: (+44) 650 3893

Fax: (+44) 651 3971

e-mail: [angus.macbeth@ed.ac.uk](mailto:angus.macbeth@ed.ac.uk)

## Introduction

Early childhood is a critical period of development which is highly susceptible to a multitude of environmental risk factors (Walker et al., 2011). Sensitive and attuned parenting has been shown to have a significant positive influence over the extent to which such risks impact child outcomes (Lengua, Honorado, & Bush, 2007; Trentacosta et al., 2008). This is particularly important in the case of families exposed to multiple co-occurring contextual risk factors e.g. low socioeconomic status, psychiatric disorder and parental experience of trauma (Masten et al., 1995), who are at higher risk of future problems (France, Freiberg, & Homel, 2010; Lanza, Rhoades, Nix, & Greenberg, 2010; Rutter, 1979). Consequently, there has been significant growth in the development of group-based parenting programmes that aim to improve parenting skills during early stages of child development (Barlow, Bergman, Kornør, Wei, & Bennett, 2016).

Parenting interventions such as ‘Incredible Years’ (Webster-Stratton & Herbert, 1994) and ‘Triple P’ (Sanders & Dadds, 1993) are effective in improving child behavioural outcomes in school-aged and pre-school children (Dretzke et al., 2009; NICE, 2006). However, evidence supporting group-based parenting interventions for parent-child dyads with children under two years old remains particularly limited (Barlow et al., 2016; Jones, Erjavec, Viktor, & Hutchings, 2016) despite public health initiatives to support families through improved parenting skills in the first 2 years of children’s lives (Allen, 2011; Commons, 2015). Methodological limitations of the current evidence base include lack of consistent reporting of drop-out rates, analysis of data from ‘completers’ only (Barlow, Smailagic, Ferriter, Bennett, & Jones, 2010), inclusion of self-selected parents (Jones et al., 2016; Wilson et al., 2012), lack of sufficient analysis power, and a failure of authors to declare financial conflicts of interest (Coyne & Kwakkenbos, 2013; Eisner, Humphreys, Wilson, & Gardner, 2015; Wilson et al., 2012). Several reviews and meta-analyses suggest that children of disadvantaged parents show poorer intervention outcomes (Lundahl, Risser, & Lovejoy, 2006; Reyno & McGrath, 2006). Furthermore, few studies have considered maternal mental health as a primary outcome (Barlow, Smailagic, Huband, Roloff, & Bennett, 2014), despite evidence for the detrimental impact of maternal depressive symptoms on early child development and quality of parenting (Walker et al., 2011). This is of importance given that measurement of change in parental mental health may act as a barometer for the effect of the intervention on the child further downstream.

From an implementation science perspective, widely evaluated and implemented parenting programmes have been limited by poor engagement of families who are at high risk of adverse developmental outcomes (Utting, Monteiro, & Ghate, 2007). It is also unclear whether specific programmes may be more successful in groups with higher levels of risk factors (Gardner, Hutchings, Bywater, & Whitaker, 2010). Thus, the evidence-base and implementation of parenting programmes for families with additional social/health needs substantially lags behind more general implementation (Evans, Davies, Williams, & Hutchings, 2015; Olds, Sadler, & Kitzman, 2007).

In this context, Mellow Parenting represents a suite of parenting programmes targeted for vulnerable families of children from the antenatal period up to age five – with Mellow Babies (MB) targeting infants aged 0-18 months and Mellow Toddlers (MT) for children aged 19 months through to pre-school (aged 4-5 years). MP programmes are multi-modal, guided by attachment theory, social learning and cognitive behaviour-theory, and diverge from other evaluated programmes through their explicit targeting of parents who often have trouble engaging in services due to complex needs such as child protection concerns, severe depression, anxiety disorders, or domestic violence (Department of Education, 2013). The evidence-base for the efficacy of MP is increasing, with a meta-analysis suggesting moderate effect sizes on parental mental health and child outcomes (MacBeth et al., 2015). These are supported by qualitative studies of its effectiveness (Birtwell, Hammond, & Puckering, 2015; Puckering et al., 2011). However, little is known about the effectiveness of MP programmes in routine practice.

The current study aimed to investigate the association between MP participation (either MT or MB) and improvements in maternal and child outcomes. Our primary focus/interest was maternal mental health, (depression, anxiety and irritability). Secondary outcomes included parenting confidence, parenting daily stress and child problem behaviour. We hypothesized that MP would be associated with positive improvements in all four outcomes. We also conducted exploratory analyses to assess whether any baseline characteristics were predictive of change across any outcomes.

## Methods

### Design

The study used an uncontrolled, prospective cohort design. Pre and post intervention questionnaires were routinely collected by Mellow Parenting (MP) group facilitators from groups occurring across the UK between August 2015 and May 2017. All participants provided written consent prior to completing any assessment and data was pseudo-anonymised. The study was approved by the University of Edinburgh Ethics Committee.

### Participants

To be eligible for MB or MT, mothers must a) have at least one child aged between birth and five years (0-18 months for MB; 19+ months for MT), b) have contact with the index child during the intervention, c) be willing to consent to take part in video feedback, and d) where possible, have their child attend the group for a lunch time activity. Additional criteria for inclusion in our analyses were a) availability of demographic data, and b) availability of data from one or more outcome measure at one or more point in time (i.e. pre or post intervention, or both). A total of n=214 mothers completed at least one session of an MP group (across 38 groups). Of those, n=183 fulfilled the criteria for inclusion in our analysis (see Figure 1), of which n= 85 were from MB, n=78 from MT, and 20 were from mixed (MB & MT) groups.

Location of intervention sites from which participants were drawn included Scotland (59.0%), England (30.1%) and Northern Ireland 10.9% (see Supplementary Table 1).

### Procedure

Intervention: MB and MT are group-based day programmes, run one day per week over 14 weeks. Each MB and MT group is run by two facilitators. The structure and delivery of both programmes are largely comparable, with groups differing only with respect to their afternoon parenting workshop topics (which are tailored for either babies or toddlers respectively, see [www.mellowparenting.org](http://www.mellowparenting.org) for more information). For each group, pre intervention measures were collected during initial home visits or group information/welcome sessions, prior to session 1 of the intervention. MP facilitators typically collect post intervention measures following sessions 13 or 14 (in case of drop-out prior to the last session), or through a home visit after the last session.

## Measures

### Maternal Mental Health & Wellbeing

The Adult Wellbeing Scale (AWS; Snaith, Constantopoulos, Jardine, & McGuffin, 1978) is a clinical scale for the self-assessment of mental health and wellbeing in four areas: Depression (5 items), Anxiety (5 items), and Outwardly- and Inwardly-Directed Irritability (4 items each). It has been used as a screening tool for postnatal mood disorder (Snaith, 1993). The 18 items are rated on a 4-point Likert scale ranging from 'yes, definitely' to 'no, not at all'. Thresholds for the indication of a possible problem are: 4-6 for Depression, 6-8 for Anxiety, 5-7 for Outwards-Directed Irritability, and 4-6 for Inwardly-Directed Irritability. Total scores range from 0 to 54 (higher scores imply greater difficulty). The internal reliability of the AWS is acceptable (Cronbach's  $\alpha = .70$  to  $.92$ ; Snaith et al., 1978). For the purposes of this paper, outcomes relating to AWS will be referred to as 'maternal mental health'.

### Parenting Confidence

The Karitane Parenting Confidence Scale (KPCS; Crnec, Barnett, & Matthey, 2008) measures perceived parental self-efficacy (PPSE) through a 15-item questionnaire. It was developed for clinical screening of parenting difficulties for parents of children aged 0-12 months. KPCS scores range from 0 to 45, with higher scores representing higher parenting confidence. The cut-off score for clinically low parenting confidence is  $<39$  (Crnec et al., 2008). KPCS has been shown to have good internal consistency (Cronbach's  $\alpha = .81$ ) and test-retest reliability ( $r = .88$ ) (Crnec et al., 2008). These data were only available for dyads where child was  $<12$  months of age.

### Parenting Daily Hassles

The Parenting Daily Hassles Scale (PDH; Crnic & Greenberg, 1990) measures 'daily parenting stress' through parental perceptions about the minor daily hassles and inconveniences associated with parenting. Parents rate the frequency and intensity/impact of 20 potential daily parenting hassles. As the PDH is primarily validated for use with toddlers, data were only available for dyads where child was  $>12$  months of age. Two summary scores are obtained pertaining to the Frequency and Intensity of parenting hassles. The Frequency subscale is rated on a 4-point Likert scale from 'rarely' to 'constantly', with scores ranging from 0 to 80 (scores above 50 indicate a high frequency of hassling happenings). The Intensity subscale is rated on a 5-point Likert scale from 'no hassle' to 'big hassle', with scores ranging from 0 to 100 (scores above 70 indicate a high degree of parenting pressure).

The internal reliability of the PDH is good, with Cronbach's alpha of .81 for Frequency and .90 for Intensity (Crnic & Greenberg, 1990).

### Child Psychosocial Behaviour

The Strengths and Difficulties Questionnaire (SDQ; Goodman, 1997) is a 25-item, parent-reported questionnaire pertaining to child competency and problem behaviours. The present study utilised the of preschool informant version of the SDQ (Youth in Mind, 2014). Items relate to the frequency of positive and negative behaviours and are rated on a 3-point Likert scale. It contains subscales relating to four areas of child difficulty: Emotional Symptoms, Peer Problems, Hyperactivity-Inattention, and Conduct Problems, in addition to a separate Prosocial Behaviour scale. The four difficulty subscales range in score from 0 to 10 and can be summed to produce a Total Difficulties score from 0 to 40 (a higher score represents greater difficulties). A higher score on the Prosocial Behaviour scale indicates greater prosocial behaviour. Scores for individual subscales and the Total Difficulties measure can be categorized into normal, borderline and abnormal ranges (Youth in Mind, 2014). A recent study utilizing a cohort of 2-year-olds showed good internal reliability for Total Difficulties (Cronbach's alpha = .84), and acceptable reliability for most subscales (Cronbach's alpha = .71 to .75), with the exception of Peer Problems which showed poor reliability (Cronbach's alpha = .54) (D'Souza et al., 2016). Due to the lack of validated data on the SDQ for infants under the age of two years, data were only available for dyads where the index child was >24 months of age.

Mothers also completed one post-intervention follow-up questionnaire of intervention satisfaction, parent-child connectedness, and help-seeking confidence. Facilitators also reported participant demographics, attendance, and one open-ended qualitative further service-uptake question via a post-intervention debrief with each mother.

### Statistical Analyses

All participants were included in the analysis, regardless of attendance. Descriptive statistics were calculated for attendance, demographics, and post-intervention follow-up questionnaires. Mean outcome scores at baseline were compared to normative means. We excluded KPCS data from N=5 mothers of toddlers who were incorrectly administered the KPCS, and due to validity concerns we excluded n=24 SDQ reports from mothers with a

recorded child age of less than 2 years old. AWS scores were combined from data for mothers who attended MB and MT or mixed groups were combined. All other outcomes were analysed individually. For demographic covariates, any categories with <5 cases were removed. We used both intention-to-treat (ITT) and “completer” analysis. For ITT analyses, linear mixed effects models were constructed using pre- and post-intervention measures as the dependant variable. Mixed effects models were chosen as they superior to other ITT approaches such as ‘last observation carried forward’ (Baraldi & Enders, 2010; Blankers, Koeter, & Schippers, 2010). The impact of the intervention on pre-post score change was examined by including a main effect of time in each model. Since demographic data was missing for several mothers, we limited the inclusion of covariates to those that showed significant main effects, or interactions with time. Lastly, the addition of a random intercept (to account for baseline differences in outcome measures) improved model fit across all outcomes (as measured via likelihood ratio tests) and was included in all final models.

For completer analyses, pre-post change in each outcome measure was analysed using paired t-tests, or Wilcoxon signed rank tests (in cases of non-normality). The sample size for AWS, KPCS, and PDH was larger in the completer analysis than the ITT analysis, as the latter involved the inclusion of covariates, which necessitated the exclusion of mothers with missing demographic data. Within-group Cohen’s *d* effect sizes were calculated and the standard deviation were corrected for dependence between responses with the equation  $SD_{corrected} = SD_{pooled}[2(1 - \rho)]^{1/2}$ , where  $\rho$  is the correlation between pre and post scores, as per (Morris, 2008). To rule out the possibility that dropout (or missing data) could have influenced our estimates of Cohen’s *d*, we used multiple imputation (MI) to impute missing values and re-calculated Cohen’s *d* for each outcome using the fully imputed dataset (Enders, 2017). MI used linear regression and the Fully Conditional Specification method with 5 imputed datasets generated for each outcome measure (Rubin, 2009; 2012) and pooled analysis results over these 5 datasets. Standard deviation (SD) was calculated as the mean SD over the 5 imputed datasets.

To investigate whether change in AWS score differed between groups, a one-way ANCOVA was performed with group (babies, toddlers or mixed) as the explanatory variable, and post-AWS score as the dependent variable (controlling for AWS at baseline). Finally, we used multiple regression to assess whether individual differences at baseline predicted attendance.



Explanatory variables included all demographic variables, whilst the dependant variable included number of sessions attended (min = 4, max = 14). Intercorrelation among predictors (multicollinearity) and presence of outliers was assessed prior to running any regressions. All statistical analyses were conducted in SPSS v24.0. All tests are two-tailed with a significance level of 0.05 unless otherwise stated. Intention for further service-uptake qualitative responses were coded with respect to two main categories (group versus individual service type) and five subcategories (see Supplementary Table 3).

## Results

### Baseline Characteristics

Table 1 shows the baseline characteristics of all mothers with demographic data, and all mothers with complete AWS data (AWS being the only outcome measure spanning all groups). Age ranged from 17-50 years old. Mean child age was 15.78 months (SD 13.69 months), with 92% of children in the age range of 0-3 years. The majority of mothers were unemployed with benefits and had a high school education or higher. Approximately 60% of mothers had a history of mental health problems. Baseline mean AWS indicated anxiety and depression were within the borderline problem range, whereas outward- and inward-directed irritability were within the normal range (Table 3). Mean KPCS scores were clinically low and skewed towards lower parenting confidence, while mean SDQ ('total difficulties') scores were 'slightly raised' to 'high', predominantly driven by Conduct Problems (Table 3). Prosocial Behaviour scores were 'slightly low', though all other subscales were in the normal range (see Table 3). Mean PDH Frequency scores were below the cut-off for indicating significant parenting stress.

### Primary outcome measures

#### *Intention-to-treat analysis*

Linear mixed effects models revealed a significant main effect of time on total AWS scores ( $\beta = -3.07$  [95% CI -4.61 to -1.53 to], SE = 0.77,  $p < 0.001$ , [n = 99]) and KPCS scores ( $\beta = 2.48$  [95% CI 1.13 to 3.86], SE = 0.68,  $p = 0.001$ , [n = 61]) in the predicted directions. In addition, all AWS subscales revealed a significant main effect of time except for Inward Irritability ( $p > 0.05$ ) (Table 2). By contrast, there was no main effect of time on either SDQ Total Difficulties ( $\beta = -1.01$  [95% CI -2.27 to 0.24], SE = 0.62,  $p = 0.11$ , [n = 41]), PDH

Frequency ( $\beta = -1.05$  [95% CI -3.08 to 0.97], SE = 1.00,  $p = 0.30$  [ $n = 50$ ]) or PDH Intensity ( $\beta = -0.79$  [95% CI -4.15 to 2.58], SE = 1.66,  $p = 0.63$ , [ $n = 50$ ]). Examining individual subscales of the SDQ revealed a significant main effect of time on Conduct Problems ( $\beta = -0.76$  [95% CI -1.44 to -0.07], SE = 0.34,  $p = 0.032$ , [ $n = 41$ ]), but none of the other subscales (all  $p > 0.05$ ).

### *Completer analysis*

Paired-samples t-tests (using complete case data only) supported the findings of the ITT analyses, with significant pre-post intervention improvements on AWS ( $t(156) = 4.749$ ,  $p < 0.001$ ,  $d=0.40$ ) and KPCS ( $t(70) = 3.483$ ,  $p = 0.001$ ,  $d = 0.43$ ) scores. All AWS subscales showed a significant change over time except Inward Irritability ( $p > 0.05$ ). As baseline KPCS scores were non-normal (Wilks-Shapiro test  $p < 0.001$ ), we also performed a Wilcoxon signed-rank test, which confirmed a significant pre-post increase in confidence ( $Z = 3.70$ ,  $p < 0.001$ ). There was no significant change in SDQ Total Difficulties ( $p > 0.05$ ,  $d=0.27$ ) or SDQ subscales demonstrated significant change over time (all  $p > 0.05$ ). PDH scores did not exhibit significant change over time, consistent with the ITT analysis. Results from the completer analysis remained similar when using MI to impute pre-or post outcome scores missing in the data except for the SDQ Conduct Problems subscale which was statistically significant using multiple imputation ( $t(184) = 2.106$ ,  $p = 0.037$ , [ $n = 41$ ]), compared to a non-significant trend in the completer analysis ( $p = 0.055$ ).

As a sensitivity analysis we conducted a one-way ANCOVA to confirm that change in AWS would not depend on programme variant. This confirmed that post AWS scores did not significantly differ between groups (babies, toddlers or mixed) ( $F(2, 156) = 2.59$ ,  $p > 0.05$ , [ $n = 157$ ]).

### **Covariates**

Several covariates included within the ITT mixed effects models were significant. For AWS, there was a main effect of maternal age ( $\beta = 0.50$  [95% CI 0.29 to 0.72], SE = 0.11,  $p < 0.001$ ) and post-intervention parent-child connection ( $\beta = -4.19$  [95% CI -6.61 to -1.76], SE = 1.22,  $p = 0.001$ ), though no significant interaction with time. A relationship status x time interaction emerged, with co-habiting mothers demonstrating greater benefit relative to single mothers ( $\beta = -3.36$  [95% -6.32 to -0.39], SE = 1.50,  $p = 0.027$ ). However, this was not significant when controlling for mother's age and post-intervention parent-child connection.

There was also a main effect of mother's age ( $\beta = -0.20$  [95% CI -0.31 to -0.07], SE = 0.07,  $p = 0.004$ ) and post-intervention parent-child connection ( $\beta = 1.98$  [95% CI 0.54 to 3.41], SE = 0.72,  $p = 0.008$ ) on KPCS scores, with no significant interaction with time. A relationship status x time interaction emerged, with married mothers demonstrating greater benefit over time relative to single mothers ( $\beta = -4.16$  [95% CI -8.23 to -0.09], SE = 2.04,  $p = 0.045$ ). This effect remained significant when controlling for mother's age and post-intervention parent-child connection ( $\beta = -4.78$  [95% CI -9.37 to -0.19], SE = 2.29,  $p = 0.041$ ). The number of sessions attended did not predict score change across any of the outcome measures (all  $p > 0.05$ ).

### **Attendance**

The mean number of sessions attended across all participants was 10.64 (SD = 2.75, [N=130]). For completers, the mean number of sessions attended was 11.07 (SD = 2.13, [N=118]). Multiple regression revealed that those with a history of mental health difficulties had higher attendance ( $\beta = 1.37$  [95% CI 0.54 to 2.20], SE = 0.42,  $p = 0.001$  [N=78]). In addition, younger mothers attended more sessions ( $\beta = -0.09$  [95% CI -0.03 to -0.15], SE = 0.03,  $p = 0.006$  [N=78]), and those who were married attended more sessions than those who were single ( $\beta = 1.78$  [95% CI 0.40 to 3.17], SE = 0.70,  $p = 0.01$  [N=78]). Neither employment nor education were significant predictors of number of sessions attended within this sample (all  $p > 0.05$ ).

### **Satisfaction and further service uptake**

Post-intervention data on participant's subjective experience of MP was available for  $n=121$  respondents (see Supplementary Table 2). Of those mothers that completed the intervention, 92.6% ( $n=112$ ) strongly agreed with the statement "I enjoyed taking part in the Mellow group", 74.4% ( $n=90$ ) strongly agreed with the statement "I feel more connected with my child after taking part in this group", and 64.5% ( $n=78$ ) strongly agreed with the statement "I feel confident in asking for help should I need it". Data from facilitators on uptake of further service was available for  $n=61$  participants (see Supplementary Table 3). Facilitators indicated that 34% ( $n=55$ ) of mothers who engaged until the end of the intervention expressed an intention to engage in further group-based community services and 3% were referred to specialised one-to-one family support ( $n=5$ ).

## Discussion

This is the first study to evaluate the impact of MP in a ‘real-life’ context and suggests that parenting programmes may be effective even for families with higher levels of risk factors. Our study demonstrates that participation in MP was associated with significantly improved maternal mental health, parenting confidence, and child conduct problems post-intervention. These results were robust across ITT and completer analyses. Although previous studies have demonstrated positive effects on child outcomes following group-based parenting programmes, few have evaluated both maternal and child outcomes, or targeted at-risk families of very young children. Participant demographics indicated multiple indices of developmental adversity among mothers, suggesting that MP is effective at targeting at-risk families. Attendance and intervention satisfaction were high amongst completers, **suggesting from an implementation science perspective that MP is acceptable to mother-child dyads with complex needs.**

Our findings are consistent with a recent study showing improvements in standardised measures of parental mental wellbeing and confidence, in vulnerable parent-child dyads who attended the Incredible Years Parent and Baby programme (Evans et al., 2015). However, parenting intervention studies have generally reported data from completers only, or have recruited participants from a single geographical region (Evans et al., 2015). The present study also builds upon existing evidence of the effectiveness of MP (MacBeth et al., 2015) by using a national UK sample of mothers routinely recruited into MP and incorporating an ITT analysis.

We also note that the improvements in maternal mental health and parenting confidence are consistent with approaches that suggest that parenting interventions should focus on both parent and child wellbeing (Alvarez, Meltzer-Brody, Mandel, & Beeber, 2015), particularly when families are targeted on the basis of parental risk factors. Moreover, high parenting confidence may act as a buffer against factors such as parental depression, stress, relationship difficulties, and compromised child development, though the direction of influence is unclear (Coleman & Karraker, 2003; T. L. Jones & Prinz, 2005). Our findings suggest that future studies of maternal characteristics in parenting interventions should aim to model the relationship between parenting confidence and mental health, as opposed to focusing on the former or latter in isolation.

Contrary to our hypothesis, MP was not associated with an overall improvement in SDQ ‘total difficulties’. This appears at odds with meta-analytic evidence for effectiveness of parenting interventions targeting at-risk parents of infants aged 0-12 months (Rayce, Rasmussen, Klest, Patras, & Pontoppidan, 2017), and existing evidence for MP being associated with reduced childhood problems. Further, in their meta-analysis of 8 MP studies, MacBeth and colleagues (2015) reported medium-level effect sizes, but cautioned against the inclusion of small, underpowered studies. This disparity may relate to differences between the present study and previous studies with regards to intervention type, broader versus narrower banding of child age, or research methodologies (Barlow et al., 2016). Furthermore, the primary focus of MP is relationship and emotional development, not behavior management.

In addition, in the present study, the conduct problems subscale of SDQ showed significant post-intervention improvement following MI analysis. This may have been driven by the fact that conduct problem scores indicated greater difficulties relative to other SDQ subscales at baseline. The finding that MP participation was associated with significant improvement in child conduct problems is encouraging in the context of the target population, as children who experience social adversity are particularly vulnerable to conduct disordered behaviour and/or mental health problems (Webster-Stratton, 1998).

However, it should be noted that the majority of studies (including the present study) assess outcomes during the intervention period or immediately post-intervention, with few reporting follow-up data (Rayce et al., 2017). **One recent MP service evaluation showed that maternal well-being (AWS; depression, anxiety and irritability) continued to improve at 6-month follow-up and was still elevated above baseline at 12-month follow-up (McGowan & McParland, 2017). Despite showing longer-term promise, these MP findings and others are limited by small samples, uncontrolled designs and lack of standardized child outcomes.** Cost-benefit research suggests that early and effective intervention for the most vulnerable children will generate the greatest financial payback in the long-term (Cunha, Heckman, Lochner, & Masterov, 2005). Thus, conclusions based on data gathered immediately post-intervention may be insufficient to inform the true effectiveness of the intervention, particularly for infants and toddlers who are experiencing a rapid period of development.

Future MP studies should incorporate follow-up assessments in order to examine the long-term effects of early interventions on child outcomes.

Finally, MP participation was not associated with improvements in ‘daily parenting stress’. It should be noted that baseline daily parenting stress scores were below the cut-off for indicating either a high frequency of hassling events, or significant parenting stress (c.f. (Crnic & Greenberg, 1990), which may have limited opportunity for further improvement. Previous studies have indicated that minor parenting hassles may influence parent-child relationships and contribute to dysfunction in families (Crnic & Greenberg, 1990). Families facing multiple contextual stressors may require support to directly reduce the family’s stress burden (Trentacosta et al., 2008). For example, child-care assistance from social supports may be necessitated to maintain parental sensitivity and enjoyment of their children, as opposed to emotional support only (Crnic & Greenberg, 1990). Caregivers’ perception of parenting stress may be a worthy indicator to examine in future research of parenting interventions, particularly for families with multiple contextual risk factors.

### **Additional Findings**

Our data suggested no differential impact of higher rates of attendance on outcome. One possible explanation is that attendance may be a poor proxy of engagement in parenting group interventions. For MP, some parents are initially referred by primarily statutory sector services and are therefore less likely to be originally attending based on their own volition, despite all parents receiving a full explanation of the programme and its objectives from MP practitioners prior to session one of the intervention. Vulnerable parents may underestimate the need for parenting support (Bussing et al., 2003), or may lack the motivation to engage with interventions (DiClemente & Velasquez, 2002). This may also reflect the reality that the relationship between service engagement and outcomes remains poorly understood, as most intervention studies do not typically consider attendance when assessing outcomes. Future studies should aim to better quantify engagement (for example through measures of in-session participation) as an alternative to physical attendance.

Exploratory analyses revealed that younger mothers, and those with a history of mental health problems, attended more intervention sessions. These findings are surprising, as engaging at-risk families into ongoing services is a known challenge (Lessing, 2013; Muzik

et al., 2015). Teenage or young adult mothers and their children are known to be at increased risk of poor health and social outcomes (Hovdestad, Shields, Williams, & Tonmyr, 2015), and the presence of mental health difficulties in mothers is associated with an increased risk of psychopathology in children (Bouvette-Turcot, Bernier, & Leblanc, 2017). Our findings also suggest MP may also be effectively utilising components associated with improved engagement and acceptability, such as content, process and access (Hutchings et al., 2007).

We also found that mothers with a partner attended more sessions, and showed greater improvements in mental health and confidence, compared to single mothers. From a public health perspective, population studies suggest that children in low-SES or single-parent families have poorer emotional and behavioural problems than their counterparts (Flouri, Midouhas, & Ruddy, 2016). In addition, single-parent status has been associated with lower retention in other parenting programmes (Orrell-Valente et al., 1999). In the present study this may highlight the role of wider family functioning in self-appraised parental competence (Knauth, 2000), and that partner support may be a protective or stabilising factor for mothers attending MP. However, the number of married mothers included in our analysis was small, thus future studies should aim to corroborate this exploratory finding in larger samples. The finding also highlights the need to better engage with vulnerable single-parent families.

## **Limitations**

The present study has several limitations. First, data were collected in a ‘real world’ setting, therefore lacks a control group for comparison. We therefore cannot assume that findings were directly caused by the MP intervention. Future studies should aim to replicate our results with the inclusion of a control group.

Second, our study relied on routine data which contained a degree of missing cases. We were unable to delineate data lost due to drop-out versus administrative non-imputation.

Furthermore, given the real-world setting we note MP facilitators may not always retain data from participants who drop out at very early stages of the intervention. Although we mitigated the impact of missing data via ITT analyses we may not have captured data from very early drop-outs. This highlights the importance of accurate data tracking and recording within health and social care settings. We also note the conflict of interest (CoI) of two authors of our study who were directly affiliated with the MP organisation. Routine MP data

is collected directly by facilitators and then collated by MP, who provided access to the data. However, these authors played no role in the analysis or initial [delete?] interpretation of the study findings, thus mitigating any risk of bias. It is important for future studies to consider and declare the impact of CoI on findings.

Third, demographic data was largely unavailable for mothers that attended less than 4 sessions. Thus, analyses exploring the relationship between SES and attendance (or change in outcome scores) were likely underpowered. In addition, demographic data such as education and employment may only partially capture SES, and future studies may wish to consider more precise measures.

Fourth, the MP database is limited to data gathered immediately pre- and post-intervention, with no follow-up data. Thus, it remains unclear whether the benefits reported in the present study are maintained over time. This may be particularly important for changes in child outcomes, which may be more likely to emerge over time when one considers that early childhood is a period of rapid development.

Fifth, although we found statistically significant improvements, our study does not shed light on the clinical significance of these improvements (Matthey, 1998), or on the mediators that drive them.

Finally, the present study used self-reported data, with child outcome data limited to dyads where child was >24 months of age due to the lack of validated data on the SDQ for younger children. Future intervention evaluations should utilize direct objective measures to assess child outcomes rather than relying solely on indirect observation (Aspland & Gardner, 2003), and consider the evaluation of young infants.

## References

- Allen, G. (2011). *Early Intervention: The next steps*. Cabinet Office, London: Crown Copyright.
- Alvarez, S. L., Meltzer-Brody, S., Mandel, M., & Beeber, L. (2015). Maternal Depression and Early Intervention: A Call for an Integration of Services. *Infants and Young Children*, 28(1), 72–87.  
<http://doi.org/10.1097/IYC.0000000000000024>



- Aspland, H., & Gardner, F. (2003). Observational Measures of Parent-Child Interaction: An Introductory Review. *Child and Adolescent Mental Health*, 8(3), 136–143. <http://doi.org/10.1111/1475-3588.00061>
- Baraldi, A. N., & Enders, C. K. (2010). An introduction to modern missing data analyses. *Journal of School Psychology*, 48(1), 5–37. <http://doi.org/10.1016/j.jsp.2009.10.001>
- Barlow, J., Bergman, H., Kornør, H., Wei, Y., & Bennett, C. (2016). Group-based parent training programmes for improving emotional and behavioural adjustment in young children. *The Cochrane Database of Systematic Reviews*, 42(8), CD003680. <http://doi.org/10.1002/14651858.CD003680.pub3>
- Barlow, J., Smailagic, N., Ferriter, M., Bennett, C., & Jones, H. (2010). Group-based parent-training programmes for improving emotional and behavioural adjustment in children from birth to three years old. *The Cochrane Database of Systematic Reviews*, 42(3), CD003680. <http://doi.org/10.1002/14651858.CD003680.pub2>
- Barlow, J., Smailagic, N., Huband, N., Roloff, V., & Bennett, C. (2014). Group-based parent training programmes for improving parental psychosocial health. *Cochrane Database of Systematic Reviews*, 9(1), 66. <http://doi.org/10.1002/14651858.CD002020.pub4>
- Birtwell, B., Hammond, L., & Puckering, C. (2015). “Me and my bump”: an interpretative phenomenological analysis of the experiences of pregnancy for vulnerable women. *Clinical Child Psychology and Psychiatry*, 20(2), 218–238. <http://doi.org/10.1177/1359104513506427>
- Blankers, M., Koeter, M. W. J., & Schippers, G. M. (2010). Missing data approaches in eHealth research: simulation study and a tutorial for nonmathematically inclined researchers. *Journal of Medical Internet Research*, 12(5), e54. <http://doi.org/10.2196/jmir.1448>
- Bouvette-Turcot, A.-A., Bernier, A., & Leblanc, É. (2017). Maternal Psychosocial Maladjustment and Child Internalizing Symptoms: Investigating the Modulating Role of Maternal Sensitivity. *Journal of Abnormal Child Psychology*, 45(1), 157–170. <http://doi.org/10.1007/s10802-016-0154-8>
- Bussing, R., Zima, B. T., Gary, F. A., Mason, D. M., Leon, C. E., Sinha, K., & Garvan, C. W. (2003). Social networks, caregiver strain, and utilization of mental health services among elementary school students at high risk for ADHD. *Journal of the American Academy of Child and Adolescent Psychiatry*, 42(7), 842–850. <http://doi.org/10.1097/01.CHI.0000046876.27264.BF>
- Coleman, P. K., & Karraker, K. H. (2003). Maternal self-efficacy beliefs, competence in parenting, and toddlers' behavior and developmental status. *Infant Mental Health Journal*, 24(2), 126–148. <http://doi.org/10.1002/imhj.10048>

- Commons, H. M. (2015) *Hansard Debate 17 December* (Vol. 603). Retrieved from <http://www.publications.parliament.uk/pa/>
- Coyne, J. C., & Kwakkenbos, L. (2013). Triple P-Positive Parenting programs: the folly of basing social policy on underpowered flawed studies. *BMC Medicine*, 11(1), 11. <http://doi.org/10.1186/1741-7015-11-11>
- Crnec, R., Barnett, B., & Matthey, S. (2008). Development of an instrument to assess perceived self-efficacy in the parents of infants. *Research in Nursing & Health*, 31(5), 442–453. <http://doi.org/10.1002/nur.20271>
- Crnic, K. A., & Greenberg, M. T. (1990). Minor parenting stresses with young children. *Child Development*, 61(5), 1628–1637.
- Cunha, F., Heckman, J., Lochner, L., & Masterov, D. (2005). *Interpreting the Evidence on Life Cycle Skill Formation*. Cambridge, MA: National Bureau of Economic Research.
- Department of Education. (2013). Find a Parenting Programme. Retrieved January 13, 2018, from <http://webarchive.nationalarchives.gov.uk/20140311170415/http://education.gov.uk/commissioning-toolkit/Programme/Detail/44>
- DiClemente, C. C., & Velasquez, M. (2002). *Motivational Interviewing*. London: The Guildford Press
- Dretzke, J., Davenport, C., Frew, E., Barlow, J., Stewart-Brown, S., Bayliss, S., et al. (2009). The clinical effectiveness of different parenting programmes for children with conduct problems: a systematic review of randomised controlled trials. *Child and Adolescent Psychiatry and Mental Health*, 3(1), 7. <http://doi.org/10.1186/1753-2000-3-7>
- Eisner, M., Humphreys, D. K., Wilson, P., & Gardner, F. (2015). Disclosure of Financial Conflicts of Interests in Interventions to Improve Child Psychosocial Health: A Cross-Sectional Study. *PloS One*, 10(11), e0142803. <http://doi.org/10.1371/journal.pone.0142803>
- Enders, C. K. (2017). Multiple imputation as a flexible tool for missing data handling in clinical research. *Behaviour Research and Therapy*, 98, 4–18. <http://doi.org/10.1016/j.brat.2016.11.008>
- Evans, S., Davies, S., Williams, M., & Hutchings, J. (2015). Short-term benefits from the Incredible Years Parents and Babies Programme in Powys. *Community Practitioner : the Journal of the Community Practitioners " & Health Visitors " Association*, 88(9), 46–48.
- Flouri, E., Midouhas, E., & Ruddy, A. (2016). Socio-economic status and family structure differences in early trajectories of child adjustment: Individual and neighbourhood effects. *Health & Place*, 37, 8–15. <http://doi.org/10.1016/j.healthplace.2015.11.005>
- France, A., Freiberg, K., & Homel, R. (2010). Beyond Risk Factors: Towards a Holistic Prevention Paradigm for Children and Young People. *British Journal of Social Work*, 40(4), 1192–1210. <http://doi.org/10.1093/bjsw/bcq010>

- Gardner, F., Hutchings, J., Bywater, T., & Whitaker, C. (2010). Who Benefits and How Does It Work? Moderators and Mediators of Outcome in an Effectiveness Trial of a Parenting Intervention. *Journal of Clinical Child & Adolescent Psychology*, 39(4), 568–580.  
<http://doi.org/10.1080/15374416.2010.486315>
- Goodman, R. (1997). The Strengths and Difficulties Questionnaire: a research note. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 38(5), 581–586.
- Hovdestad, W., Shields, M., Williams, G., & Tonmyr, L. (2015). Vulnerability within families headed by teen and young adult mothers investigated by child welfare services in Canada. *Health Promotion and Chronic Disease Prevention in Canada : Research, Policy and Practice*, 35(8-9), 143–150.
- Hutchings, J., Bywater, T., Daley, D., Gardner, F., Whitaker, C., Jones, K., et al. (2007). Parenting intervention in Sure Start services for children at risk of developing conduct disorder: pragmatic randomised controlled trial. *Bmj*, 334(7595), 678. <http://doi.org/10.1136/bmj.39126.620799.55>
- Jones, C. H., Erjavec, M., Viktor, S., & Hutchings, J. (2016). Outcomes of a Comparison Study into a Group-Based Infant Parenting Programme. *Journal of Child and Family Studies*, 25(11), 3309–3321.  
<http://doi.org/10.1007/s10826-016-0489-3>
- Jones, T. L., & Prinz, R. J. (2005). Potential roles of parental self-efficacy in parent and child adjustment: a review. *Clinical Psychology Review*, 25(3), 341–363. <http://doi.org/10.1016/j.cpr.2004.12.004>
- Knauth, D. G. (2000). Predictors of parental sense of competence for the couple during the transition to parenthood. *Research in Nursing & Health*, 23(6), 496–509. [http://doi.org/10.1002/1098-240X\(200012\)23:6<496::AID-NUR8>3.0.CO;2-1](http://doi.org/10.1002/1098-240X(200012)23:6<496::AID-NUR8>3.0.CO;2-1)
- Lanza, S. T., Rhoades, B. L., Nix, R. L., & Greenberg, M. T. (2010). Modeling the interplay of multilevel risk factors for future academic and behavior problems: A person-centered approach. *Development and Psychopathology*, 22(02), 313. <http://doi.org/10.1017/S0954579410000088>
- Lengua, L. J., Honorado, E., & Bush, N. R. (2007). Contextual risk and parenting as predictors of effortful control and social competence in preschool children. *Journal of Applied Developmental Psychology*, 28(1), 40–55. <http://doi.org/10.1016/j.appdev.2006.10.001>
- Lessing, L. A. (2013). Serving Highly Vulnerable Families in Home-Visitation Programs. *Infant Mental Health Journal*, 34(5), 376–390.  
<http://doi.org/10.1002/imhj.21399>
- Lundahl, B., Risser, H. J., & Lovejoy, M. C. (2006). A meta-analysis of parent training: moderators and follow-up effects. *Clinical Psychology Review*, 26(1), 86–104. <http://doi.org/10.1016/j.cpr.2005.07.004>
- MacBeth, A., Law, J., McGowan, I., Norrie, J., Thompson, L., & Wilson, P. (2015). Mellow Parenting: systematic review and meta-analysis of an

- intervention to promote sensitive parenting. *Developmental Medicine & Child Neurology*, 57(12), 1119–1128. <http://doi.org/10.1111/dmcn.12864>
- Masten, A. S., Coatsworth, J. D., Neemann, J., Gest, S. D., Tellegen, A., & Garmezy, N. (1995). The structure and coherence of competence from childhood through adolescence. *Child Development*, 66(6), 1635–1659.
- Matthey, S. (1998).  $p < .05$  - But is it Clinically Significant? Practical Examples for Clinicians. *Behaviour Change*, 15(2), 140–146. <http://doi.org/10.1017/S0813483900003247>
- McGowan, I. W., & McParland, D. (2017). I'm not just a Mummy. An evaluation of the Mellow Parenting Programme in the Southern Health and Social Care Trust. Southern Health & Social Care Trust. Armagh
- Morris, S. B. (2008). Estimating Effect Sizes From Pretest-Posttest-Control Group Designs. *Organizational Research Methods*, 11(2), 364–386. <http://doi.org/10.1177/1094428106291059>
- Muzik, M., Rosenblum, K. L., Alfafara, E. A., Schuster, M. M., Miller, N. M., Waddell, R. M., & Kohler, E. S. (2015). Mom Power: preliminary outcomes of a group intervention to improve mental health and parenting among high-risk mothers. *Archives of Women's Mental Health*, 18(3), 507–521. <http://doi.org/10.1007/s00737-014-0490-z>
- NICE. (2006). National Institute for Health and Clinical Excellence. Parent training/education programmes in the management of children with conduct disorders. London.
- Olds, D. L., Sadler, L., & Kitzman, H. (2007). Programs for parents of infants and toddlers: recent evidence from randomized trials. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 48(3-4), 355–391. <http://doi.org/10.1111/j.1469-7610.2006.01702.x>
- Orrell-Valente, J. K., Valente, E., Jr., Laird, R. D., Bierman, K. L., Coie, J. D., Dodge, K. A., et al. (1999). If It's Offered, Will They Come? Influences on Parents' Participation in a Community-Based Conduct Problems Prevention Program. *American Journal of Community Psychology*, 27(6), 753–783. <http://doi.org/10.1023/A:1022258525075>
- Puckering, C., Connolly, B., Werner, C., Toms-Whittle, L., Thompson, L., Lennox, J., & Minnis, H. (2011). Rebuilding relationships: a pilot study of the effectiveness of the Mellow Parenting Programme for children with Reactive Attachment Disorder. *Clinical Child Psychology and Psychiatry*, 16(1), 73–87. <http://doi.org/10.1177/1359104510365195>
- Rayce, S. B., Rasmussen, I. S., Klest, S. K., Patras, J., & Pontoppidan, M. (2017). Effects of parenting interventions for at-risk parents with infants: a systematic review and meta-analyses. *BMJ Open*, 7(12), e015707. <http://doi.org/10.1136/bmjopen-2016-015707>
- Reyno, S. M., & McGrath, P. J. (2006). Predictors of parent training efficacy for child externalizing behavior problems--a meta-analytic review. *Journal of*

- Child Psychology and Psychiatry, and Allied Disciplines*, 47(1), 99–111.  
<http://doi.org/10.1111/j.1469-7610.2005.01544.x>
- Rubin, D. B. (2009). Multiple Imputation for Nonresponse in Surveys. (D. B. Rubin, Ed.). Hoboken, NJ, USA: John Wiley & Sons.  
<http://doi.org/10.1002/9780470316696>
- Rubin, D. B. (2012). Multiple Imputation after 18+ Years. *Journal of the American Statistical Association*, 91(434), 473–489.  
<http://doi.org/10.1080/01621459.1996.10476908>
- Rutter, M. (1979). Maternal deprivation, 1972-1978: new findings, new concepts, new approaches. *Child Development*, 50(2), 283–305.
- Sanders, M. R., & Dadds, M. R. (1993). Behavioral Family Intervention. Allyn & Bacon.
- Snaith, P. (1993). Measuring anxiety and depression. *The Practitioner*, 237(1528), 554–559.
- Snaith, P., Constantopoulos, A. A., Jardine, M. Y., & McGuffin, P. (1978). A clinical scale for the self-assessment of irritability. *The British Journal of Psychiatry : the Journal of Mental Science*, 132, 164–171.
- Trentacosta, C. J., Hyde, L. W., Shaw, D. S., Dishion, T. J., Gardner, F., & Wilson, M. (2008). The relations among cumulative risk, parenting, and behavior problems during early childhood. *Journal of Child Psychology and Psychiatry*, 40(11), ???–??? <http://doi.org/10.1111/j.1469-7610.2008.01941.x>
- Utting, D., Monteiro, H., & Ghate, D. (2007). Interventions for children at risk of developing antisocial personality disorder. London, UK: Policy Research Bureau.
- Walker, S. P., Wachs, T. D., Grantham-McGregor, S., Black, M. M., Nelson, C. A., Huffman, S. L., et al. (2011). Inequality in early childhood: risk and protective factors for early child development. *Lancet (London, England)*, 378(9799), 1325–1338. [http://doi.org/10.1016/S0140-6736\(11\)60555-2](http://doi.org/10.1016/S0140-6736(11)60555-2)
- Webster-Stratton, C. (1998). Preventing conduct problems in Head Start children: strengthening parenting competencies. *Journal of Consulting and Clinical Psychology*, 66(5), 715–730.
- Webster-Stratton, C., & Herbert, M. (1994). Troubled families - Problem Children. Chichester, England: Wiley and Sons.
- Wilson, P., Rush, R., Hussey, S., Puckering, C., Sim, F., Allely, C. S., et al. (2012). How evidence-based is an “evidence-based parenting program?” A PRISMA systematic review and meta-analysis of Triple P. *BMC Medicine*, 10(1), 130. <http://doi.org/10.1186/1741-7015-10-130>
- Youth in Mind. (2014). Retrieved January 13, 2018, from <http://www.sdqinfo.com/norms/UK3yearNorm.html>

Table 1: Baseline demographics for total sample and completers analysis.

Study Variable	All mothers	Completers N = 103
Age (n [%])	<b>N = 142</b>	
17-24	54 (38.0)	39 (37.9)
25-29	47 (33.1)	29 (28.2)
30-39	33 (23.2)	28 (27.2)
40-50	8 (5.6)	7 (6.8)
History of mental health problems (n [%])	<b>N = 143</b>	
Yes	88 (61.5)	65 (63.1)
No	55 (38.5)	38 (36.9)
Employment Status (n [%])	<b>N = 148</b>	
Full-time employment	6 (4.1)	6 (5.8)
Part-time employment	19 (12.8)	14 (13.6)
Unemployed – no benefits	14 (9.5)	8 (7.8)
Unemployed – job seekers allowance	4 (2.7)	3 (2.9)
Unemployed – with benefits	105 (70.9)	72 (69.9)
Education Status (n [%])	<b>N = 128</b>	
Still at school	1 (0.8)	1 (1.0)
Didn't finish school	21 (16.4)	17 (16.5)
Finished school (Standard grades/GCSEs)	54 (42.2)	45 (43.7)
College (NC/A levels or higher)	52 (40.6)	40 (38.8)
Relationship Status (n [%])	<b>N = 147</b>	
Single	58 (39.5)	40 (38.8)
Married/civil partnership	16 (10.9)	11 (10.7)
Co-habiting	47 (32.0)	34 (33.0)
In a relationship, not co-habiting	24 (16.3)	17 (16.5)
Separated	1 (0.7)	1 (0.9)
Divorced	1 (0.7)	0 (0)

Notes: Total sample represents all mothers with one or more category of demographic data available at baseline. Completers represents all mothers who completed the Adult Wellbeing Scale (AWS) both pre-and post-intervention (and with full demographic data available).

Table 2: Beta coefficient for a main effect of time on all primary outcome measures and subscales via mixed effects modelling.

Outcomes Measure	<i>N</i>	<i>β</i>	<i>95% CI</i>	<i>SE</i>	<i>p</i>
AWS Total	99	-3.07	-4.61 to -1.53	0.77	0.000**
AWS Depression	99	-0.95	-1.50 to -0.39	0.28	0.001**
AWS Anxiety	99	-1.15	-1.70 to -0.60	0.27	0.000**
AWS Outward Irritability	99	-0.48	-0.90 to -0.07	0.21	0.023*
AWS Inward Irritability	99	-0.26	-0.76 to 0.23	0.25	0.295
KPCS	61	2.48	1.13 to 3.86	0.68	0.001**
SDQ Total Difficulties	41	-1.01	-2.27 to 0.24	0.62	0.11
SDQ Emotional Symptoms	41	-0.08	-0.79 to 0.64	0.35	0.83
SDQ Conduct Problems	41	-0.76	-1.44 to -0.07	0.34	0.032*
SDQ Hyperactivity/Inattention	41	0.14	-0.54 to 0.82	0.34	0.68
SDQ Peer Relation Problems	41	-0.14	-0.73 to 0.45	0.29	0.64
SDQ Prosocial Behaviour	41	0.27	-0.40 to 0.95	0.33	0.42
PDH Frequency	50	-1.05	-3.08 to 0.97	1.00	0.30
PDH Intensity	50	-0.79	-4.15 to 2.58	1.66	0.63

Notes: AWS: Adult Wellbeing Scale; KPCS: Karitane Parenting Confidence Scale; PDHS: Parenting Daily Hassles Scale; SDQ: Strengths and Difficulties Questionnaire. In the case of AWS, KPCS, and PDHS, sample sizes differ to the completers and MI analyses due to demographic data being included as covariates). Mothers with missing demographic data being omitted from the mixed effects analyses.  $\beta$  = model coefficient for a main effect of time on the dependant variable, 95% CI = Confidence interval for  $\beta$ , SE = standard error. \*\*  $p \leq 0.001$ , \* $p < 0.05$ .

Table 3: Mean pre-and post- scores for all primary outcome measures and subscales for the completer sample and imputed dataset

Outcomes		Pre mean (SD)	Post mean (SD)	<i>t</i>	<i>d</i> (95% CI)
<b>AWS Total</b>	complete, n = 157	20.38 (9.50)	17.55 (8.79)	4.749	0.40 ** (0.17 to 0.62)
	imputed, n = 179	20.45 (9.77)	17.83 (9.07)	4.487	0.36 ** (0.10 to 0.52)
AWS Depression	complete, n = 157	6.07 (2.92)	5.17 (2.58)	4.261	0.41 ** (0.13 to 0.55)
	imputed, n = 179	6.09 (2.89)	5.22 (2.61)	4.296	0.33** (0.10 to 0.52)
AWS Anxiety	complete, n = 157	6.96 (3.21)	5.99 (3.21)	4.207	0.31 ** (0.09 to 0.54)
	imputed, n = 179	6.88 (3.26)	6.05 (3.25)	3.535	0.29 ** (0.08 to 0.49)
AWS Outward Irritability	complete, n = 157	3.92 (2.68)	3.42 (2.40)	3.047	0.25 * (0.46 to 0.01)
	imputed, n = 179	3.90 (2.72)	3.52 (2.50)	2.114	0.18 * (0.03 to 0.38)
AWS Inward Irritability	complete, n = 157	3.38 (3.06)	3.03 (2.70)	1.821	0.15 (0.36 to 0.08)
	imputed, n = 179	3.44 (3.14)	3.17 (3.31)	1.427	0.10 (-0.32 to 0.1)
<b>KPCS</b>	complete, n=71	38.00 (5.94)	40.23 (4.00)	3.483	0.43 ** (0.03 to 0.70)
	imputed, n=73	38.05 (5.89)	40.24 (3.98)	3.479	0.42** (0.03 to 0.69)
<b>SDQ Total</b>	complete, n=36	15.92 (5.17)	15.00 (6.29)	1.475	0.27 (0.18 to 0.75)
	imputed, n=41	16.39 (5.31)	15.42 (6.45)	1.653	0.28 (0.13 to 0.74)
SDQ Emotion	complete, n = 36	3.00 (1.91)	2.86 (2.05)	0.380	0.07 (-0.40 to 0.53)
	imputed, n = 41	2.90 (1.83)	2.84 (2.33)	0.171	0.03 (-0.40 to 0.46)
SDQ Conduct	complete, n=36	4.36 (2.24)	3.69 (2.08)	1.986	0.33 ≠ (0.14 to 0.79)
	imputed, n = 41	4.59 (2.42)	3.82 (2.16)	2.105	0.36 * (0.10 to 0.78)
SDQ Hyper	complete, n=36	5.72 (2.08)	5.92 (2.37)	0.584	0.10 (-0.36 to 0.57)
	imputed, n = 41	5.88 (2.32)	5.87 (2.48)	0.015	0.01 (-0.43 to 0.44)
SDQ Peer	complete, n=36	2.81 (1.74)	2.61 (1.70)	0.642	0.11 (-0.57 to 0.35)
	imputed, n = 41	2.71 (1.65)	2.59 (1.69)	0.414	0.07 (-0.37 to 0.5)
SDQ Prosocial	complete, n=36	6.72 (2.34)	7.00 (2.04)	0.818	0.14 (-0.33 to 0.59)
	imputed, n = 41	6.73 (1.31)	6.98 (2.03)	0.685	0.16 (-0.23 to 0.64)
<b>PDH Frequency</b>	complete, n=62	41.40 (10.82)	40.11 (9.95)	1.557	0.19 (0.16 to 0.54)
	imputed, n=73	41.47 (10.82)	40.08 (10.01)	1.517	0.21 (0.12 to 0.53)
<b>PDH Intensity</b>	complete, n=60	45.70 (12.75)	44.92 (13.36)	0.560	0.07 (0.29 to 0.43)
	imputed, n=71	45.71 (13.41)	45.30 (14.00)	0.295	0.04 (0.29 to 0.36)



Notes: AWS: Adult Wellbeing Scale; KPCS: Karitane Parenting Confidence Scale; PDHS: Parenting Daily Hassles Scale; SDQ: Strengths and Difficulties Questionnaire; complete = complete case; imputed = imputed values; d = within-group Cohen's d effect sizes; SD = Standard Deviation; t = paired-samples t-test statistic; 95% CI = 95% Confidence Interval; \*\*  $p \leq 0.001$ , \* $p < 0.05$ ,  $\neq 0.055$ .